

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14:42:42

Input Set A:\es.txt  
Output Set N:\CRF3\01162002\J028384.raw

**ENTERED**

3 <110> APPLICANT: COMPATIGENE INC.  
4 PERREAULT, Claude  
5 MCBRIDE, Kevin  
7 <120> TITLE OF INVENTION: Mammalian SIMP, gene sequence and uses thereof in cancer therapy  
9 <130> FILE FEFERENCE: 5600-74  
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/028,384  
C--> 11 <141> CURRENT FILING DATE: 2001-12-20  
11 <160> NUMBER OF SEQ ID NOS: 13  
13 <170> SOFTWARE: PatentIn version 3.1  
15 <210> SEQ ID NO: 1  
16 <211> LENGTH: 3481  
17 <213> TYPE: DNA  
18 <213> ORGANISM: Homo sapiens  
20 <220> FEATURE  
21 <221> NAME/KEY CDS  
22 <222> LOCATION: (1)..(2481)  
23 <223> OTHER INFORMATION  
24 <400> SEQUENCE 1  
25 atq gcg gaa ccc tcg gcc ccg gag agc aag cac aag tcg tcc ctc aac 48  
26 Met Ala Glu Pro Ser Ala Pro Glu Ser Lys His Lys Ser Ser Leu Asn  
27 1 5 10 15  
28 tcg tcc ccg tag agt ggc ctc atg gcc ctg gga aac agc ccg cac ggc 96  
29 Ser Ser Pro Trp Ser Gly Leu Met Ala Leu Gly Asn Ser Arg His Gly  
30 20 25 30  
31 cac cac ggu ccc ggg gcc cag tgc gcg cac aag gcg ggg ggc ggc 144  
32 His His Gly Pro Gly Ala Gln Cys Ala His Lys Ala Ala Gly Gly Ala  
33 35 40 45  
34 gcg ccg ccg aag ccg gcc ccg gcg ggg ctg tcc ggg ggg ctg tcc cag 192  
35 Ala Pro Pro Lys Pro Ala Pro Ala Gly Leu Ser Gly Gly Leu Ser Gln  
36 50 55 60  
37 ccg gct gga ttt cag tcg ctt ctc tcc acc atc ctc ttc ctg gcc 240  
38 Pro Ala Gly Trp Gln Ser Leu Leu Ser Phe Thr Ile Leu Phe Leu Ala  
39 65 70 75 80  
40 tgg ctt gcc ggc ttc agc tcg cgc ctc ttc gcc gtc atc cgc ttc gaa 288  
41 Trp Leu Ala Gly Phe Ser Ser Arg Leu Phe Ala Val Ile Arg Phe Glu  
42 85 90 95  
43 agc atc atc cac gag ttc gac ccg tgg ttt aac tat aga tca aca cat 336  
44 Ser Ile Ile His Glu Phe Asp Pro Trp Phe Asn Tyr Arg Ser Thr His  
45 100 105 110  
46 cat ctt gca tct cat ggg ttc tat gaa ttt tta aat tgg ttt gat gaa 384  
47 His Leu Ala Ser His Gly Phe Tyr Glu Phe Leu Asn Trp Phe Asp Glu  
48 115 120 125  
49 aga gca tgg tat cca cta gga aga ata gta ggt ggt act gtt tac cca 432  
50 Arg Ala Trp Tyr Pro Leu Gly Arg Ile Val Gly Gly Thr Val Tyr Pro  
51 130 135 140  
52 ggg ttg atq ata acc gct ggc ctt att cat tgg att tta aat aca ttg 480  
53 Gly Leu Met Ile Thr Ala Gly Leu Ile His Trp Ile Leu Asn Thr Leu

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14 42:42

Input Set : A:\es.txt  
Output Set: N:\CRF3\01162002\J028384.raw

65	145	150	155	160	
67	aac ata act gtt cac ata aga gac gta tgt gtg ttc ctt gca cca a <sup>t</sup> t				528
68	Asn Ile Thr Val His Ile Arg Asp Val Cys Val Phe Leu Ala Pro Thr				
69	165	170	175		
71	ttt a <sup>c</sup> g ggc ctt aca tct ata tct act ttc ctg ctt aca aga gaa ctt				576
72	Phe Ser Gly Leu Thr Ser Ile Ser Thr Phe Leu Leu Thr Arg Glu Leu				
73	180	185	190		
75	tgg aac caa gga gca gga ctt tta gct gct tgt ttt att gct att gta				624
76	Trp Asn Gln Gly Ala Gly Leu Leu Ala Ala Cys Phe Ile Ala Ile Val				
77	195	200	205		
79	cca ggc tac ata tct cgg tca gta gct g <sup>a</sup> g tcc ttt gat aat gaa g <sup>c</sup> g				672
80	Pro Gly Tyr Ile Ser Arg Ser Val Ala Gly Ser Phe Asp Asn Glu Gly				
81	210	215	220		
83	att gct att ttt gca ctt cag ttc aca tac tat tta tgg gta aaa tct				720
84	Ile Ala Ile Phe Ala Leu Gln Phe Thr Tyr Tyr Leu Trp Val Lys Ser				
85	225	230	235	240	
87	gta aaa act ggg tca gti ttt tgg aca atg tgc tgc tgc tta tcc tat				768
88	Val Lys Thr Gly Ser Val Phe Trp Thr Met Cys Cys Cys Leu Ser Tyr				
89	245	250	255		
91	ttc tat atg gtc tct gct tgg ggt ggt tat gta ttt atc atc aat ctt				816
92	Phe Tyr Met Val Ser Ala Trp Gly Gly Tyr Val Phe Ile Ile Asn Leu				
93	260	265	270		
95	att cca ctc cat gta ttt gtg ttg tta ctg atg cag a <sup>g</sup> a tac agc a <sup>a</sup> a				864
96	Ile Pro Leu His Val Phe Val Leu Leu Met Gln Arg Tyr Ser Lys				
97	275	280	285		
99	aga gtc tac ata gca tai a <sup>g</sup> c act ttc tac att gtg ggt tta ata tta				912
100	Arg Val Tyr Ile Ala Tyr Ser Thr Phe Tyr Ile Val Gly Leu Ile Leu				
101	290	295	300		
103	tca atg cag ata cct ttt gtg gga ttc cag cca atc a <sup>g</sup> a aca a <sup>g</sup> t gaa				960
104	Ser Met Gln Ile Pro Phe Val Gly Phe Gln Pro Ile Arg Thr Ser Glu				
105	305	310	315	320	
107	cac atg gca gct gca ggt gtc ttt gca ttg ctg caa gct tat gct ttc				1008
108	His Met Ala Ala Ala Gly Val Phe Ala Leu Leu Gln Ala Tyr Ala Phe				
109	325	330	335		
111	ttg cag tat ctg aga gac cga tta aca aaa caa gag ttc cag acc ctt				1056
112	Leu Gln Tyr Leu Arg Asp Arg Leu Thr Lys Gln Glu Phe Gln Thr Leu				
113	340	345	350		
115	ttc ttt ttg ggt gta tca cta gct gca ggt gct gtt ttc ctt agt gtc				1104
116	Phe Phe Leu Gly Val Ser Leu Ala Ala Gly Ala Val Phe Leu Ser Val				
117	355	360	365		
119	atc tat ttg act tat aca ggt tac att gca cca ttg agt ggc agg ttt				1152
120	Ile Tyr Leu Thr Tyr Thr Gly Tyr Ile Ala Pro Trp Ser Gly Arg Phe				
121	370	375	380		
123	tat tca ttg tgg gat act ggg tat gca aaa ata cac att cca att att				1200
124	Tyr Ser Leu Trp Asp Thr Gly Tyr Ala Lys Ile His Ile Pro Ile Ile				
125	385	390	395	400	
127	gca tca gtg tct gag cat caa cct acg act tgg gtg tct ttc itc ttt				1248
128	Ala Ser Val Ser Glu His Gln Pro Thr Thr Trp Val Ser Phe Phe Phe				
129	405	410	415		

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14:42:42

Input Set : A:\es.txt  
Output Set: N:\CRF3\01162002\J028384.raw

131 gat cta cat att ctt gta tgt acc ttc cca gca ggc ctt tgg ttc tgc	1296
132 Asp Leu His Ile Leu Val Cys Thr Phe Pro Ala Gly Leu Trp Phe Cys	
133 420 425 430	
135 atc aaa aat atc aac gat gaa aya gta ttt gtt gct cta tat gca atc	1341
136 Ile Lys Asn Ile Asn Asp Glu Arg Val Phe Val Ala Leu Tyr Ala Ile	
137 435 440 445	
139 agt gtc ttc ttt gtc gtg atg gtg cja ctg atg tig act ttg	1392
140 Ser Ala Val Tyr Phe Ala Gly Val Met Val Arg Leu Met Leu Thr Leu	
141 450 455 460	
143 act cca gtc gtt tgc atg ctg tct gca att gcc ttt tca aat gtt Ctt	1440
144 Thr Pro Val Val Cys Met Leu Ser Ala Ile Ala Phe Ser Asn Val Phe	
145 465 470 475 480	
147 gag can tat ttt ggt gat gac atg aaa agg gaa aat cca cct gtg gat	1488
148 Glu His Tyr Leu Gly Asp Asp Met Lys Arg Glu Asn Pro Pro Val Glu	
149 485 490 495	
151 gac aac agt gat gaa gat gac aua aga aac caa gga aat tig tat gat	1536
152 Asp Ser Ser Asp Glu Asp Asp Lys Arg Asn Gin Gly Asn Leu Tyr Asp	
153 500 505 510	
156 aag gca ggt aai gtc agg aaa cat gca act gaa cag gaa aua act gaa	1584
156 Lys Ala Gly Lys Val Arg Lys His Ala Thr Glu Gln Glu Lys Thr Glu	
157 515 520 525	
159 gag gga tta ggc cct aat ata aua ayc att gtc acc atg tig atg ctg	1632
160 Glu Gly Leu Gly Pro Asn Ile Lys Ser Ile Val Thr Met Leu Met Leu	
161 530 535 540	
163 atg cta ttc atg atg ttt gtc cac tgc acc tgg gtc aca ayc aat	1680
164 Met Leu Leu Met Met Phe Ala Val His Cys Thr Trp Val Thr Ser Asn	
165 545 550 555 560	
167 gcc tac tct ayt cca ayt gta gtc ctg gcc tca tac aat cat gat ggc	1728
168 Ala Tyr Ser Ser Pro Ser Val Val Leu Ala Ser Tyr Asn His Asp Gly	
169 565 570 575	
171 acc agg aat atc tta gat gat ttt aga gaa gct tac ttt tgg cta agg	1776
172 Thr Arg Asn Ile Leu Asp Asp Phe Arg Glu Ala Tyr Phe Trp Leu Arg	
173 580 585 590	
175 caa aat aca gat gaa cat gca cca gta atg tct tgg tgg gat tat ggc	1824
176 Gln Asn Thr Asp Glu His Ala Arg Val Met Ser Trp Asp Tyr Gly	
177 595 600 605	
179 tat caq ata gct gga atg gct aat aga act ayc tig gtg gat aat aac	1872
180 Tyr Gln Ile Ala Gly Met Ala Asn Arg Thr Thr Leu Val Asp Asn Asn	
181 610 615 620	
183 acc tgg aat aac ayc cac ata gca ctg gtg gga aua gct atg tct tct	1920
184 Thr Trp Asn Asn Ser His Ile Ala Leu Val Gly Lys Ala Met Ser Ser	
185 625 630 635 640	
187 aat gaa aca gca gcc tat aaa atc atg agg act cta gat gta gat tat	1968
188 Asn Glu Thr Ala Ala Tyr Lys Ile Met Arg Thr Leu Asp Val Asp Tyr	
189 645 650 655	
191 gtt ttg gtt att ttt gga ggg gtt att ggc tat tct ggt gat gat atc	2016
192 Val Leu Val Ile Phe Gly Gly Val Ile Gly Tyr Ser Gly Asp Asp Ile	
193 660 665 670	
195 aac aaa ttt ctc tgg atg gtt agg ata gct gaa gga cat ccc aaa	2064

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14 42:42

Input Set : A:\es.txt  
Output Set: N:\CRF3\01162002\J028384.raw

196	Asn	Lys	Phe	Leu	Trp	Met	Val	Arg	Ile	Ala	Glu	Gly	Glu	His	Pro	Lys
197			675			680					685					
198	gac	att	cgg	gaa	agt	gac	tat	ttt	acc	cca	cag	gga	gaa	ttc	cgt	gta
199																2112
200	Asp	Ile	Arg	Glu	Ser	Asp	Tyr	Phe	Thr	Pro	Gln	Gly	Glu	Phe	Arg	Val
201											695		700			
202	gpc	aat	gca	gga	ttc	cct	act	ttg	ttg	aat	tgc	ttt	atg	tat	aaa	atg
203																2160
204	Asp	Lys	Ala	Gly	Ser	Pro	Thr	Leu	Leu	Asn	Cys	Leu	Met	Tyr	Lys	Met
205	705					710				715			720			
206	tca	tac	tac	aga	ttt	gga	gaa	atg	cag	ctg	gat	ttt	cgt	aca	ccc	cca
207																2208
208	Ser	Tyr	Tyr	Arg	Phe	Gly	Glu	Met	Gln	Leu	Asp	Phe	Arg	Thr	Pro	Pro
209										725		730		735		
210	ggt	ttt	gac	cga	aca	cg	aat	gtt	gag	att	gja	aat	aag	gac	att	aaa
211	Gly	Phe	Asp	Arg	Thr	Arg	Asn	Ala	Glu	Ile	Gly	Asn	Lys	Asp	Ile	Lys
212							740			745			750			
213	tcc	aaa	cac	ttt	gaa	gaa	gcc	ttt	aca	tca	gaa	cac	tgg	ttt	gtt	agg
214	Phe	Lys	His	Leu	Glu	Ala	Phe	Thr	Ser	Glu	His	Trp	Leu	Val	Arg	
215							755			760			765			
216	ata	tat	aaa	gta	aaa	gca	cct	gat	aac	agg	gag	aca	tta	gat	cac	aaa
217	Ile	Tyr	Lys	Val	Lys	Ala	Pro	Asp	Asn	Arg	Glu	Thr	Leu	Asp	His	Lys
218							770			775			780			
219	cct	aga	gtc	acc	aac	att	ttc	cca	aaa	cag	aag	tat	ttg	tca	aaq	aag
220	Pro	Arg	Val	Thr	Asn	Ile	Phe	Pro	Lys	Gln	Lys	Tyr	Leu	Ser	Lys	
221							785			790			795			2352
222	ata	cac	aaa	agg	aaa	ttt	ggt	ttt								
223	ttt															
224	ttt															
225	ttt															
226	ttt															
227	ttt															
228	ttt															
229	ttt															
230	ttt															
231	ttt															
232	ttt															
233	ttt															
234	ttt															
235	ttt															
236	ttt															
237	ttt															
238	ttt															
239	ttt															
240	ttt															
241	ttt															
242	ttt															
243	ttt															
244	ttt															
245	ttt															
246	ttt															
247	ttt															
248	ttt															
249	ttt															
250	ttt															
251	ttt															
252	ttt															
253	ttt															
254	ttt															
255	ttt															
256	ttt															
257	ttt															
258	ttt															
259	ttt															
260	ttt															
261	ttt															
262	ttt															
263	ttt															
264	ttt															
265	ttt															
266	ttt															
267	ttt															
268	ttt															
269	ttt															
270	ttt															
271	ttt															

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14:42:42

Input Set : A:\es.txt  
Output Set: N:\CRF3\01162002\J028384.raw

272	115	120	125
275	Arg Ala Trp Tyr Pro Leu Gly Arg Ile Val Gly Gly Thr Val Tyr Pro		
276	130	135	140
279	Gly Leu Met Ile Thr Ala Gly Leu Ile His Trp Ile Leu Asn Thr Leu		
280	145	150	155
283	Asn Ile Thr Val His Ile Arg Asp Val Cys Val Phe Leu Ala Pro Thr		160
284	165	170	175
287	Phe Ser Gly Leu Thr Ser Ile Ser Thr Phe Leu Leu Thr Arg Glu Leu		
288	180	185	190
291	Trp Asn Gln Gly Ala Gly Leu Leu Ala Ala Cys Phe Ile Ala Ile Val		
292	195	200	205
295	Pro Gly Tyr Ile Ser Arg Ser Val Ala Gly Ser Phe Asp Asn Glu Gly		
296	210	215	220
299	Ile Ala Ile Phe Ala Leu Gln Phe Thr Tyr Tyr Leu Trp Val Lys Ser		
300	225	230	235
303	Val Lys Thr Gly Ser Val Phe Trp Thr Met Cys Cys Leu Ser Tyr		240
304	245	250	255
307	Phe Tyr Met Val Ser Ala Trp Gly Gly Tyr Val Phe Ile Ile Asn Leu		
308	260	265	270
311	Ile Pro Leu His Val Phe Val Leu Leu Leu Met Gln Arg Tyr Ser Lys		
312	275	280	285
315	Arg Val Tyr Ile Ala Tyr Ser Thr Phe Tyr Ile Val Gly Leu Ile Leu		
316	290	295	300
319	Ser Met Gln Ile Pro Phe Val Gly Phe Gln Pro Ile Arg Thr Ser Glu		
320	305	310	315
323	His Met Ala Ala Ala Gly Val Phe Ala Leu Leu Gln Ala Tyr Ala Phe		320
324	325	330	335
327	Leu Gln Tyr Leu Arg Asp Arg Leu Thr Lys Gln Glu Phe Gln Thr Leu		
328	340	345	350
331	Phe Ile Leu Gly Val Ser Leu Ala Ala Gly Ala Val Phe Leu Ser Val		
332	355	360	365
335	Ile Tyr Leu Thr Tyr Thr Gly Tyr Ile Ala Pro Trp Ser Gly Arg Phe		
336	370	375	380
339	Tyr Ser Leu Trp Asp Thr Gly Tyr Ala Lys Ile His Ile Pro Ile Ile		
340	385	390	395
343	Ala Ser Val Ser Glu His Gln Pro Thr Thr Trp Val Ser Phe Phe Phe		400
344	405	410	415
347	Asp Leu His Ile Leu Val Cys Thr Phe Pro Ala Gly Leu Trp Phe Cys		
348	420	425	430
351	Ile Lys Asn Ile Asn Asp Glu Arg Val Phe Val Ala Leu Tyr Ala Ile		
352	435	440	445
355	Ser Ala Val Tyr Phe Ala Gly Val Met Val Arg Leu Met Leu Thr Leu		
356	450	455	460
359	Thr Pro Val Val Cys Met Leu Ser Ala Ile Ala Phe Ser Asn Val Phe		
360	465	470	475
363	Glu His Tyr Leu Gly Asp Asp Met Lys Arg Glu Asn Pro Pro Val Glu		480
364	485	490	495
367	Asp Ser Ser Asp Glu Asp Asp Lys Arg Asn Gln Gly Asn Leu Tyr Asp		
368	500	505	510

VERIFICATION SUMMARY  
PATENT APPLICATION: US/10/028,384

DATE: 01/16/2002  
TIME: 14:42:43

Input Set : A:\es.txt  
Output Set: N:\CRF3\01162002\J028384.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date